η. SALAMI

Page 1 of 7
TC 1600 MAIL ROOM
RECEIVED
AUG -9 20
1648

ENTERED

```
                                          DATE: 08/02/2000
         RAW SEQUENCE LISTING             TIME: 13:46:20
         PATENT APPLICATION:  US/09/000,004

         Input Set : A:\Sabe-137.app
         Output Set: N:\CRF3\08022000\I000004.raw

              SEQUENCE LISTING
 4  (1) GENERAL INFORMATION:
 6        (i) APPLICANT: WEBB, Elizabeth Ann
 7                       MARGETTS, Mary Brigid
 8                       COX, John Cooper
 9                       FRAZER, Ian
10                       MCMILLAN, Nigel Alan John
11                       WILLIAMS, Mark Philip
12                       MOLONEY, Margaret Bridget
13                       Holland
14                       EDWARDS, Stirling John
16        (ii) TITLE OF INVENTION: PAPILLOMAVIRUS POLYPROTEIN CONSTRUCTS
18        (iii) NUMBER OF SEQUENCES: 50
20        (iv) CORRESPONDENCE ADDRESS:
21             (A) ADDRESSEE: FOLEY & LARDNER
22             (B) STREET: 3000 K Street, N.W.
23             (C) CITY: Washington
24             (D) STATE: D.C.
25             (E) COUNTRY: U.S.A.
26             (F) ZIP: 20007-5109
28        (v) COMPUTER READABLE FORM:
29             (A) MEDIUM TYPE: Floppy disk
30             (B) COMPUTER: IBM PC compatible
31             (C) OPERATING SYSTEM: PC-DOS/MS-DOS
32             (D) SOFTWARE: PatentIn Release #1.0, Version #1.30
34        (vi) CURRENT APPLICATION DATA:
35             (A) APPLICATION NUMBER: US/09/000,004
36             (B) FILING DATE: 21-Apr-1998
37             (C) CLASSIFICATION:
43        (vii) PRIOR APPLICATION DATA:
40             (A) APPLICATION NUMBER: WO PCT/AU96/00473
41             (B) FILING DATE: 26-JUL-1996
44             (A) APPLICATION NUMBER: AU PN 4439/95
45             (B) FILING DATE: 27-JUL-1995
47        (viii) ATTORNEY/AGENT INFORMATION:
48             (A) NAME: BENT, Stephen A.
49             (B) REGISTRATION NUMBER: 29,768
50             (C) REFERENCE/DOCKET NUMBER: 017227/0137
52        (ix) TELECOMMUNICATION INFORMATION:
53             (A) TELEPHONE: (202) 672-5300
54             (B) TELEFAX: (202) 672-5399
57  (2) INFORMATION FOR SEQ ID NO: 1:
59        (i) SEQUENCE CHARACTERISTICS:
60             (A) LENGTH: 30 base pairs
61             (B) TYPE: nucleic acid
62             (C) STRANDEDNESS: single
63             (D) TOPOLOGY: linear
```

C--> 35
C--> 36

8/2/00

RAW SEQUENCE LISTING
PATENT APPLICATION: US/09/000,004

DATE: 08/02/2000
TIME: 13:46:20

Input Set : A:\Sabe-137.app
Output Set: N:\CRF3\08022000\I000004.raw

```
65      (ii) MOLECULE TYPE: DNA (genomic)                              30
70      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:
72 GCGCCCCGGG ATGGAAAGTG CAAATGCCTC
74 (2) INFORMATION FOR SEQ ID NO: 2:
76      (i) SEQUENCE CHARACTERISTICS:
77          (A) LENGTH: 42 base pairs
78          (B) TYPE: nucleic acid
79          (C) STRANDEDNESS: single
80          (D) TOPOLOGY: linear
82      (ii) MOLECULE TYPE: DNA (genomic)
87      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:                        42
89 GCGCTCTAGA CCATGGAAGC TTGGGTAACA TGTCTTCCAT GC
91 (2) INFORMATION FOR SEQ ID NO: 3:
93      (i) SEQUENCE CHARACTERISTICS:
94          (A) LENGTH: 49 base pairs
95          (B) TYPE: nucleic acid
96          (C) STRANDEDNESS: single
97          (D) TOPOLOGY: linear
99      (ii) MOLECULE TYPE: DNA (genomic)
104     (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:                        49
106 GCGCTCTAGA GAGCTCGGTA CCACTAGTGG AGCACCAAAC ATTGGGAAG
108 (2) INFORMATION FOR SEQ ID NO: 4:
110     (i) SEQUENCE CHARACTERISTICS:
111         (A) LENGTH: 31 base pairs
112         (B) TYPE: nucleic acid
113         (C) STRANDEDNESS: single
114         (D) TOPOLOGY: linear
116     (ii) MOLECULE TYPE: DNA (genomic)
121     (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4:                        31
123 GCGCAGATCT TAGGCGTAGC TGAACTGTTA C
125 (2) INFORMATION FOR SEQ ID NO: 5:
127     (i) SEQUENCE CHARACTERISTICS:
128         (A) LENGTH: 33 base pairs
129         (B) TYPE: nucleic acid
130         (C) STRANDEDNESS: single
131         (D) TOPOLOGY: linear
133     (ii) MOLECULE TYPE: DNA (genomic)
138     (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 5:                        33
140 GCGCCCATGG GAAGTGGTGC CTGTACAAAT AGC
142 (2) INFORMATION FOR SEQ ID NO: 6:
144     (i) SEQUENCE CHARACTERISTICS:
145         (A) LENGTH: 32 base pairs
146         (B) TYPE: nucleic acid
147         (C) STRANDEDNESS: single
148         (D) TOPOLOGY: linear
150     (ii) MOLECULE TYPE: DNA (genomic)
155     (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 6:                        32
157 GCGCTCTAGA TTGCTGTGTG GTAACAATAT AG
159 (2) INFORMATION FOR SEQ ID NO: 7:
```

```
161      (i) SEQUENCE CHARACTERISTICS:
162          (A) LENGTH: 34 base pairs
163          (B) TYPE: nucleic acid
164          (C) STRANDEDNESS: single
165          (D) TOPOLOGY: linear
167     (ii) MOLECULE TYPE: DNA (genomic)
172     (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 7:              34
174 GCGCAAGCTT CATGGAAGAC ATGTTACCCT AAAG
176 (2) INFORMATION FOR SEQ ID NO: 8:
178      (i) SEQUENCE CHARACTERISTICS:
179          (A) LENGTH: 30 base pairs
180          (B) TYPE: nucleic acid
181          (C) STRANDEDNESS: single
182          (D) TOPOLOGY: linear
184     (ii) MOLECULE TYPE: DNA (genomic)
189     (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 8:              30
191 GCGCCCATGG GGTCTTCGGT GCGCAGATGG
193 (2) INFORMATION FOR SEQ ID NO: 9:
195      (i) SEQUENCE CHARACTERISTICS:
196          (A) LENGTH: 35 base pairs
197          (B) TYPE: nucleic acid
198          (C) STRANDEDNESS: single
199          (D) TOPOLOGY: linear
201     (ii) MOLECULE TYPE: DNA (genomic)
206     (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 9:              35
208 GCGCGAGCTC GCGGACGATT CAGGTACAGA AAATG
210 (2) INFORMATION FOR SEQ ID NO: 10:
212      (i) SEQUENCE CHARACTERISTICS:
213          (A) LENGTH: 33 base pairs
214          (B) TYPE: nucleic acid
215          (C) STRANDEDNESS: single
216          (D) TOPOLOGY: linear
218     (ii) MOLECULE TYPE: DNA (genomic)
223     (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 10:             33
225 GCGCGGTACC TAAAGTTCTA ACAACTGTTC CTG
227 (2) INFORMATION FOR SEQ ID NO: 11:
229      (i) SEQUENCE CHARACTERISTICS:
230          (A) LENGTH: 32 base pairs
231          (B) TYPE: nucleic acid
232          (C) STRANDEDNESS: single
233          (D) TOPOLOGY: linear
235     (ii) MOLECULE TYPE: DNA (genomic)
240     (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 11:             32
242 GCGCGGTACC GAAGCAATAG CCAAGCGTTT AG
244 (2) INFORMATION FOR SEQ ID NO: 12:
246      (i) SEQUENCE CHARACTERISTICS:
247          (A) LENGTH: 33 base pairs
248          (B) TYPE: nucleic acid
249          (C) STRANDEDNESS: single
```

```
250          (D) TOPOLOGY: linear
252     (ii) MOLECULE TYPE: DNA (genomic)
257     (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 12:                    33
259 GCGCACTAGT CAATAGGTGC AGTGACATAA ATC
261 (2) INFORMATION FOR SEQ ID NO: 13:
263     (i) SEQUENCE CHARACTERISTICS:
264          (A) LENGTH: 32 base pairs
265          (B) TYPE: nucleic acid
266          (C) STRANDEDNESS: single
267          (D) TOPOLOGY: linear
269     (ii) MOLECULE TYPE: DNA (genomic)
274     (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 13:                    32
276 GCGCTCTAGA CTAACATGTC AATTTAATGA TG
278 (2) INFORMATION FOR SEQ ID NO: 14:
280     (i) SEQUENCE CHARACTERISTICS:
281          (A) LENGTH: 31 base pairs
282          (B) TYPE: nucleic acid
283          (C) STRANDEDNESS: single
284          (D) TOPOLOGY: linear
286     (ii) MOLECULE TYPE: DNA (genomic)
291     (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 14:                    31
293 GCGCGAGCTC ATTCATATAT ATATAATCAC C
295 (2) INFORMATION FOR SEQ ID NO: 15:
297     (i) SEQUENCE CHARACTERISTICS:
298          (A) LENGTH: 30 base pairs
299          (B) TYPE: nucleic acid
300          (C) STRANDEDNESS: single
301          (D) TOPOLOGY: linear
303     (ii) MOLECULE TYPE: DNA (genomic)
308     (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 15:                    30
310 GCGCCCCGGG ATGGAAGCAA TAGCCAAGCG
312 (2) INFORMATION FOR SEQ ID NO: 16:
314     (i) SEQUENCE CHARACTERISTICS:
315          (A) LENGTH: 51 base pairs
316          (B) TYPE: nucleic acid
317          (C) STRANDEDNESS: single
318          (D) TOPOLOGY: linear
320     (ii) MOLECULE TYPE: DNA (genomic)
325     (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 16:                    51
327 GCGCTCTAGA CCATGGGGTA CCGAGCTCCA ATAGGTGCAG TGACATAAAT C
329 (2) INFORMATION FOR SEQ ID NO: 17:
331     (i) SEQUENCE CHARACTERISTICS:
332          (A) LENGTH: 32 base pairs
333          (B) TYPE: nucleic acid
334          (C) STRANDEDNESS: single
335          (D) TOPOLOGY: linear
337     (ii) MOLECULE TYPE: DNA (genomic)
342     (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 17:                    32
344 GCGCTCTAGA CTAACATGTC AATTTAATGA TG
```

```
346 (2) INFORMATION FOR SEQ ID NO: 18:
348       (i) SEQUENCE CHARACTERISTICS:
349             (A) LENGTH: 36 base pairs
350             (B) TYPE: nucleic acid
351             (C) STRANDEDNESS: single
352             (D) TOPOLOGY: linear
354      (ii) MOLECULE TYPE: DNA (genomic)                          36
359      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 18:
361 GCGCAGATCT CTCGAGATTC ATATATATAT AATCAC
363 (2) INFORMATION FOR SEQ ID NO: 19:
365       (i) SEQUENCE CHARACTERISTICS:
366             (A) LENGTH: 1107 base pairs
367             (B) TYPE: nucleic acid
368             (C) STRANDEDNESS: single
369             (D) TOPOLOGY: linear
371      (ii) MOLECULE TYPE: DNA (genomic)
374      (ix) FEATURE:
375             (A) NAME/KEY: CDS
376             (B) LOCATION: 1..1104
379      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 19:                   48
381 ATG GAA AGT GCA AAT GCC TCC ACG TCT GCA ACG ACC ATA GAC CAG TTG
382 Met Glu Ser Ala Asn Ala Ser Thr Ser Ala Thr Thr Ile Asp Gln Leu
383  1               5                  10                  15
385 TGC AAG ACG TTT AAT CTA TCT ATG CAT ACG TTG CAA ATT AAT TGT GTG   96
386 Cys Lys Thr Phe Asn Leu Ser Met His Thr Leu Gln Ile Asn Cys Val
387                  20                  25                  30
389 TTT TGC AAG AAT GCA CTG ACC ACA GCA GAG ATT TAT TCA TAT GCA TAT  144
390 Phe Cys Lys Asn Ala Leu Thr Thr Ala Glu Ile Tyr Ser Tyr Ala Tyr
391             35                  40                  45
393 AAA CAC CTA AAG GTC CTG TTT CGA GGC GGC TAT CCA TAT GCA GCC TGC  192
394 Lys His Leu Lys Val Leu Phe Arg Gly Gly Tyr Pro Tyr Ala Ala Cys
395         50                  55                  60
397 GCG TGC TGC CTA GAA TTT CAT GGA AAA ATA AAC CAA TAT AGA CAC TTT  240
398 Ala Cys Cys Leu Glu Phe His Gly Lys Ile Asn Gln Tyr Arg His Phe
399 65                  70                  75                  80
401 GAT TAT GCT GGA TAT GCA ACA ACA GTT GAA GAA GAA ACT AAA CAA GAC  288
402 Asp Tyr Ala Gly Tyr Ala Thr Thr Val Glu Glu Glu Thr Lys Gln Asp
403                  85                  90                  95
405 ATC TTA GAC GTG CTA ATT CGG TGC TAC CTG TGT CAC AAA CCG CTG TGT  336
406 Ile Leu Asp Val Leu Ile Arg Cys Tyr Leu Cys His Lys Pro Leu Cys
407             100                 105                 110
409 GAA GTA GAA AAG GTA AAA CAT ATA CTA ACC AAG GCG CGG TTC ATA AAG  384
410 Glu Val Glu Lys Val Lys His Ile Leu Thr Lys Ala Arg Phe Ile Lys
411         115                 120                 125
413 CTA AAT TGT ACG TGG AAG GGT CGC TGC CTA CAC TGC TGG ACA ACA TGC  432
414 Leu Asn Cys Thr Trp Lys Gly Arg Cys Leu His Cys Trp Thr Thr Cys
415     130                 135                 140
417 ATG GAA GAC ATG TTA CCC AAG CTT CCA TGG GAA GTG GTG CCT GTA CAA  480
418 Met Glu Asp Met Leu Pro Lys Leu Pro Trp Glu Val Val Pro Val Gln
```

VERIFICATION SUMMARY
PATENT APPLICATION: US/09/000,004

DATE: 08/02/2000
TIME: 13:46:21

Input Set : A:\Sabe-137.app
Output Set: N:\CRF3\08022000\I000004.raw

L:35 M:220 C: Keyword misspelled or invalid format, [(A) APPLICATION NUMBER:]
L:36 M:220 C: Keyword misspelled or invalid format, [(B) FILING DATE:]